

Using the Fun Toolkit and Other Survey Methods to Gather Opinions in Child Computer Interaction

Janet C Read

Child Computer Interaction Group
University of Central Lancs, UK
jcread@uclan.ac.uk

Stuart MacFarlane

Child Computer Interaction Group
University of Central Lancs, UK
sjmacfarlane@uclan.ac.uk

ABSTRACT

The paper begins with a review of some of the current literature on the use of survey methods with children. It then presents four known concerns with using survey methods for opinion gathering and reflects on how these concerns may impact on studies in Child Computer Interaction. The paper then investigates the use of survey methods in Child Computer Interaction and investigates the Fun Toolkit. Three new research studies into the efficacy and usefulness of the tools are presented and these culminate in some guidelines for the future use of the Fun Toolkit. The authors then offer some more general guidelines for HCI researchers and developers intending to use survey methods in their studies with children. The paper closes with some thoughts about the use of survey methods in this interesting but complex area.

Keywords

Survey methods, Questionnaires, Interviews, Fun Toolkit, Children, Evaluation

ACM Classification Keywords

H5.2 User Interfaces, Evaluation / Methodology

INTRODUCTION

The method of eliciting information by questioning is commonly referred to as a survey method. Surveys are a long established instrument for gathering opinions and information from people and they are often used in HCI to gather opinions about products as well as to identify requirements for products. In a recent study into the use of methods with HCI practitioners in the Nordic community, survey methods were highlighted as being especially useful [2].

The term survey has many meanings but for the purposes of this paper, survey methods are defined as questionnaires,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IDC '06, June 7-9, 2006 Tampere, Finland
Copyright 2006 ACM 1-59593-316-6/06/07... \$5.00

rating scales and structured interviews [10]. Thus, free discussion and free form reporting is not especially considered. The main contribution of this paper is to offer a clearer understanding about the usefulness of some of the tools in the Fun Toolkit and to present guidelines for survey methods for children.

RELATED RESEARCH

As early as the 1890's surveys have been reported as being used with children [3]. However, research about the efficacy of the different methods of surveying children is relatively scarce, and in particular, when children are asked to contribute opinions,, studies that examine the validity and reliability of the children's responses are rare [5].

Why Ask Children?

In the field of Child Computer Interaction it is common to find studies that report the use of survey methods with children. In some of these studies, children are asked to contribute ideas and suggestions for future or partially completed designs. Examples include the use of surveys to elicit detail about the mental models that children have [25], or their use to gather requirements for interfaces [26]. More commonly, surveys are used in evaluation studies, where children are asked to comment on the appeal or usefulness of a product or supply some sort of product rating [24].

There are several valid reasons for asking children for their opinions of interactive products. One is that adults and children live in different worlds and for that reason adults may not understand what children want, "*Survey researchers are realising that information on children's opinions, attitudes and behaviour should be collected directly from the children; proxy-reporting is no longer considered good enough.*" [4]. Secondly, there is a move to include children in decisions about their own environments; this has arisen from a greater awareness that children are actors and participants rather than onlookers in society. "*In most of the western world, it is now recognised that children have a voice that should be heard and there is a new demand for research that focuses on children as actors in their own right.*" [7]. A third reason

for talking to children about their interactive technologies, perhaps for some people the most motivating, is that involving children in the design and evaluation of their own artefacts is fun and rewarding for researchers, developers and, more importantly, for children [19].

What Can Go Wrong?

As outlined earlier, surveys methods rely on the use of a question and answer process. Asking good questions is not easy, and for some children, understanding and interpreting the question, and formulating an appropriate response can be very difficult. Understanding the question answer process can assist researchers in designing good surveys [5]. Breakwell, [8], describes four stages in a question-answer process:

1. Understanding and interpreting the question being asked.
2. Retrieving the relevant information from memory.
3. Integrating this information into a summarised judgement.
4. Reporting this judgement by translating it to the format of the presented response scale.

Factors that impact on question answering include developmental effects including language ability, reading age, and motor skills, as well as temperamental effects such as confidence, self-belief and the desire to please.

Research into the completion of surveys has revealed four major concerns that are important in understanding how children respond to surveys and therefore important to consider in the design of surveys. The first two, Satisficing and Optimising and Suggestibility, are phenomena that will have an impact on the design of survey studies. The second two, language effects and question formats, are rather more concerned with the detailed design of the question and answer processes.

Satisficing And Optimising

Satisficing theory identifies two processes that explain some of the differences in the reliability of responses, especially in surveys where respondents are being asked to pass attitudinal judgments [14]. For research validity, optimising is the preferred process; this occurs when a survey respondent goes thoughtfully and carefully through all four stages of the question and answer sequence. Satisficing is the opposite approach and occurs when a respondent gives more or less superficial responses that generally appear reasonable or acceptable, but without having gone through all the steps involved in the question-answer process.

The degree or level of satisficing is known to be related to the motivation of the respondent, the difficulties of the task, and the cognitive abilities of the respondent [4]. It appears obvious therefore, that if a child misunderstands a

question or finds it difficult to answer then the child is susceptible to 'satisfice'.

Suggestibility

Suggestibility is particularly important with relation to survey research with children, because it "concerns the degree to which children's encoding, storage, retrieval and reporting of events can be influenced by a range of social and psychological factors." [30]. In any survey, the interviewer or researcher has an effect as even when the interviewer is trying hard not to impact on the question answer process, when the respondents are children it is sometimes impossible to not intervene. In one study it was reported "there were many silences that needed some input if only to make the children less uncomfortable." [26].

Even where there is no deliberate intervention the interviewer has an effect. In one study it was shown that children are likely to give different responses depending on the status of the interviewer. This was illustrated when a research assistant pretending to be a police officer asked children questions about their experience with a babysitter. The children then assumed that the nature of the experience was bad and thus the interviews yielded inaccurate and misleading results [32]. It seems that authority figures may inevitably yield different results, as the child may want to please the person administering the survey [9].

The gender and age of the interviewer or person conducting the survey can also have an effect on the reliability or detail of responses provided by children. Borgers et al, (2004) provide an example stating: "There is anecdotal evidence from surveys on drugs in Germany that teenagers were far more open to elderly female interviewers and not to the young or youngish interviewers." [5].

Specific Question Formats

The way in which children are asked questions in surveys has an impact on the reliability of the response. Breakwell et al, (1995) report that "There is a strong acquiescence response bias in children: children tend to say 'yes', irrespective of the question or what they think about it." [8]. In one study with 5-year-old children there were several inaccuracies in questions that relied on the yes/no format [9].

Free-recall questions have been shown to be useful with children, especially in spoken surveys. One study involved children who had experience of being treated in an emergency room for an injury. A few days later, children were interviewed with free recall question formats such as "Tell me what happened" and specific questions like "Where did you hurt yourself?" both being used. It was shown that as the questions became more specific i.e. "Did you hurt your knee?" the response reliability decreased [18]. One problem for the researcher with free-recall answers is in coding the responses. Several studies use this

method but often the papers omit the detail about how the information was then coded [34], [28].

One widely used question format is the use of Visual Analogue Scales (VAS). A VAS uses pictorial representations that children use to identify their feelings or opinions. This approach has been adopted as an alternative to the traditional open-ended and closed question formats although some researchers suggest that VAS can only be used with children aged around seven and over [31]. Studies in Child Computer Interaction have shown them to be useful for younger children, but have also noted that when these scales are used to elicit opinions about software or hardware products, younger children are inclined to almost always indicate the highest score on the scale [21].

Below are two examples of Visual Analogue Scales developed for children for different purposes.

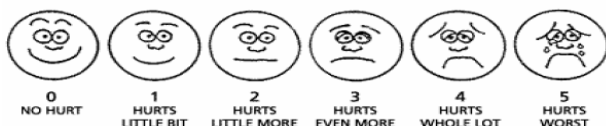


Figure 1 - Wong Baker pain rating scale [1]

Language Effects

Children have varying abilities in spoken and written language and this makes the design of questions for surveys problematic. Research suggests that language in surveys is especially important and that vague and ambiguous words should be avoided [4]. When visual analogue scales or multi-choice responses are used, it is advised that the response options should be completely labelled to help children to produce more reliable responses [6].

Children are known to take things literally and the way they understand words cannot always be predicted; in one study it was noted that when a group of children were asked if they had been on a 'school field trip' they replied 'no' because they did not refer to the trip as a 'school field trip.' [13]. In a more recent study, it was noted that when children were asked how good they thought a writing activity had been, some children gave an opinion of their writing as a product, thus interpreting the question in a completely unexpected way [27].

THE USE OF SURVEY METHODS IN CCI

Within the community, researchers in Child Computer Interaction use several survey methods. These include very simple Yes/No methods like 'Did you like it', the use of more structured question and answer methods, and the use of toolkits.

Much of the work on surveys with children has been carried out by Hanna and Ridsden. They developed the first funometer [11], and more recently reported a study into the usefulness of several rating methods [12]. This study suggested several areas for further research, in particular it reflected on the possibilities for pairwise comparisons for usability testing of products. In line with work by other researchers, the study concluded that by and large, children had high opinions of the products that they evaluated.

Airey et al., presented work with quite young children that used tangible devices to record rankings. The children found the method easy to use, but again, as in all these studies, the authors were cautious about reading too much into the findings [1].

Another influence on survey methods with children has been the Fun Toolkit [22]. This has been well used with the ideas and the tools that were introduced in the early paper being used in several studies. These studies have included papers that have evaluated gaming applications [16], [17], multimedia [29] and educational applications [15].

Many studies in CCI rely on the use of survey methods to provide usability information. There is, therefore, a need to determine how useful these methods are.

The Fun Toolkit Revisited

In its original form, the Fun Toolkit comprised four special tools, a Smileyometer, a Funometer, an Again - Again Table, and a Fun Sorter and also supported the idea of measuring remembrance and of using video footage to score engagement. The Smileyometer is shown in Figure 2 and is a discrete Likert type scale.



Figure 2 - A Smileyometer

This was used before an experience, to measure expectations, and after an experience to apply a judgment score. The Funometer is very similar to the Smileyometer but uses a continuous scale; this is based on the one in [11]. Early work found that the Funometer and the Smileyometer were essentially similar and so the Funometer has seldom been used since and is not discussed further here. The two other special tools, the Fun Sorter and the Again - Again table both measured different things. The Fun Sorter, seen in Figure 3, allowed children to rank items against one or more constructs. This was intended to record the children's opinions of the technology or activity, to gain a measure of the child's engagement.

Name of child.....Age.....Boy / Girl







| | | | | |
|----------|---|---|---|---|
| | Best | | | Worst |
| Most fun |  |  |  |  |

Figure 3 - A Completed Fun Sorter

The Again - Again table (Figure 4) was designed to capture an idea of engagement by asking the children whether or not they would do the activity again.

Would you like to do it again?

| | Yes | Maybe | No |
|--|-----|-------|----|
|  clock | ✓ | | |
|  dive | | ✓ | |
| | | | |

Name of child.....Age.....Boy / Girl

Figure 4 - A Completed Again - Again table

The early study of the Fun Toolkit described the theoretical basis for the tools and reported three studies in which the tools were evaluated [22]. The first study used sixteen children aged between 6 and 10, the second used 45 children aged 7 and 8, and the third used 53 children aged 8 – 10. These studies determined that the ‘Fun measures’ were easy for the children to use and that there was some correlation across tools. The three major findings from these early studies were that younger children tended to score most things as ‘Brilliant’ on the Smileyometer, that children demonstrated a desire to ‘play fair’¹ on the Fun Sorter, and that there was little difference between how good the children expected something to be with what they eventually rated it. Also identified in this early work was that some children, especially the younger ones, had difficulties with the constructs in the Fun Sorter.

The studies that are presented here have been designed to test the Fun Toolkit for validity and to investigate some of the concerns and ideas that were proposed by the original authors..

Study 1 – The Effect Of Age On Smileyometer Results

It was reported in [22] and [15] that the Smileyometer was not a useful tool for young children as too many children

¹ Rearranging their initial orderings so that one item was not always ‘last’ or ‘first’.

tended to choose the extreme values (mostly high ones) and so the data had little variability. A study was designed to further investigate these results and to determine whether or not age had an effect on responses and on variability

To do this, 47 children aged between 7 and 9 and 26 children aged 12 and 13 attended one of three similar events at the University in 2005 and 2006. In between taking part in a number of other activities, the children were presented with a website that linked to a selection of online games suitable for their ages. The children all saw the same list of games and were allowed to explore the games, and try them out. (Sometimes it is easy to get children to participate in research!) For each game that they tried, they were asked to record their opinion of the game on a Smileyometer. Most children tried and graded several games. In total the older children completed 119 Smileyometers and the younger children completed 121.

The Smileyometers were scored from 1 to 5 where 1 represented ‘awful’ and 5 represented ‘brilliant’.

Almost half the younger children gave everything they saw a score of five, whereas the older children were more discriminating. These proportions can be seen in figures 5 and 6.

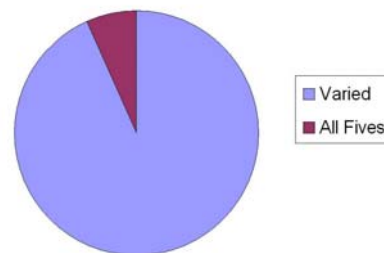


Figure 5 - Variability in the Older Children's Responses

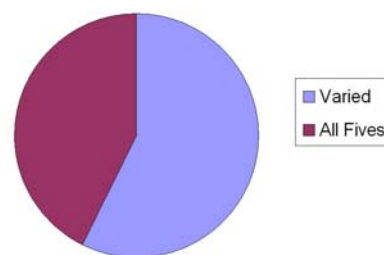


Figure 6 - Variability in the Younger Children's Responses

The mean score for the younger children was 3.860 (or nearly ‘really good’), while the mean for older children was 3.487. There is a significant difference (Mann-Whitney U = 5786.5, p (two-tailed) = 0.006) supporting the hypothesis that responses to this kind of question vary with age. Scores get lower as children get older.

Given that the younger children in this study were aged between 7 and 9, it is expected, but not confirmed here, that

even younger children would have even higher mean scores. Anecdotally we have noted that smileyometer scales are of limited value with very young children as nearly all of them pick 'brilliant', whatever the actual experience.

This work implies that the Smileyometer is more useful with older children and that, given the large number of children that give all fives; it should be used with caution with small samples of young children.

Study 2 – Sorting, Constructs And Doing It All Again

In the early work by Read et al., there was a suggestion that the Fun Sorter and the Again – Again table were related, although this was not tested in a statistical way [23]. To validate the Fun Sorter and the Again-Again table and to assess whether they are measuring the same construct in the minds of the children a study was designed that asked children to evaluate a number of interactive devices using both methods. This would then allow a check for consistency.

This work involved 15 children aged 7 and 8 in the use of three different writing interfaces; a keyboard interface, a tablet PC, and pencil-and-paper. As part of the study, the children ranked the three items for how much fun they were, and separately, for how usable they were, using two Fun Sorters. In addition, they also completed an Again-Again table (for the three items) to indicate whether they would like to use them again. All 15 children completed the evaluations.

The Fun Sorter results were coded as 3 for the highest ranked, 2 for the next, and 1 for the lowest. The Again-Again results were coded as 3 for 'yes', 2 for 'maybe' and 1 for 'no'. Non-parametric correlations were carried out.

| | Yes | Maybe | No |
|---|-----|-------|----|
| 3 | 12 | 1 | 2 |
| 2 | 4 | 8 | 3 |
| 1 | 2 | 2 | 11 |

Table 1 - Pairs of Scores for Fun and Again – Again

| | Yes | Maybe | No |
|---|-----|-------|----|
| 3 | 7 | 3 | 4 |
| 2 | 7 | 7 | 1 |
| 1 | 3 | 2 | 10 |

Table 2 - Pairs of Scores for Ease of Use and Again – Again

The results (seen in Tables 1 and 2) showed a very strong correlation (Spearman's rho = 0.526, p<0.0005) between the Again-Again scores and the Sorter results for fun.

There was a much weaker (non-significant) correlation between the Again-Again scores and the rankings for ease of use. These results are particularly interesting as it can be concluded that (1) many of the children at this age could distinguish between the constructs of 'fun' and 'ease of use', confirming results recorded in [15], and, (2) the Again-Again table is assessing the same construct as a Fun Sorter that is asking children to rank for 'fun', but a different construct to a Fun Sorter that is ranking according to 'ease of use'. This indicates that the major factor in a child's decision about whether they want to use an interactive product again is how much fun it was.

Study 3 – Doing It All Again

In the original work on the Fun Sorter, the suggestion was made that the Again – Again table measured a facet of fun related to 'returnance'. From the findings in Study 2, it appeared that the Again –Again table was measuring the same construct as the Fun Sorter with a Fun Construct, what was not clear, was how this related to the Smileyometer results.

Twenty four children aged 8 and 9 participated in an evaluation in which they rated activities using a Smileyometer and an Again – Again table.

The Again-Again results were coded as 3 for 'yes', 2 for 'maybe' and 1 for 'no', and the Smileyometers were coded from 1 to 5 where 5 was brilliant. Non-parametric correlations were carried out.

In total, 60 results were rated. The correlation between the results was very high 0.780 (Spearman's rho) p<0.0005. The 'pairs' for each tool are shown in Table 3..

| | Yes | Maybe | No |
|---|-----|-------|----|
| 5 | 29 | 2 | |
| 4 | 4 | 3 | 1 |
| 3 | 4 | 5 | |
| 2 | | 2 | 4 |
| 1 | | 1 | 5 |

Table 3 - Pairs of Scores for Smileyometer and Again - Again

This suggests that this is again the same construct (fun) that is being measured.

The Fun Toolkit Revised

From these three studies, and from the reported use of the toolkit from various sources, the authors offer the following suggestions for its future use.

The Smileyometer is an adequate tool for an easy and attractive method of scoring an opinion but is more useful with older children. There is no real point in using a Smileyometer and an Again –Again table as both measure the same construct, having said which, there have been

instances where the Again – Again table has seemed to have more validity, possibly due to the shift in the emphasis of the evaluation (it is hypothesized that in the Again – Again table, younger children are less likely to be affected by suggestibility as it seems that they are not really judging the software developer, rather just casting their own opinion [20]).

As the Fun Sorter with a construct of Fun is measuring the same construct as the Smileyometer and the Again – Again table, it perhaps makes sense to save its use for other constructs such as ‘ease of use’ or ‘easiest to learn’. The studies reported here suggest that children that are quite young can understand different constructs.

Given the tendency of children to ‘Play Fair’ it is recommended that different activities or technologies with the Smileyometer and different constructs with the Fun Sorter are presented on different pieces of paper.

DISCUSSION

The results presented here reveal that the decision whether or not to use an interactive product or not is based on how much fun it is perceived to be. This may not be especially useful in the early stages of designing a product, but it does suggest that, when comparing products, fun is a useful differential. Indeed, it is the case, that, at different points of a product lifecycle, different survey approaches need to be used and different questions asked.

It is often the case that at the point children are asked for opinions of interactive products, it is to confirm a research hypothesis or to satisfy a developer that what they have made is sound. This ‘rubber stamping’ of results can be carried out with a Yes/No question (rather flawed due to acquiescence), a Smileyometer (also rather flawed for the same reasons, but possibly worthwhile with older children) or an Again - Again table (potentially less flawed). Where there is a comparison of items, features, or products, the Fun Sorter with the construct Fun, or the Again - Again table will reveal similar results.

Where the intention is to gather opinions in order to improve or modify a product, more general survey methods are needed. It is possible to use the Again - Again table and the Fun Sorter to rank features of a product that are attractive to use, but this would only reveal areas for development rather than indicate specific improvements to be made. In these instances, a short written or verbal survey may be required.

Guidelines For Surveys With Children

There are several useful approaches that can be taken to make the surveying process valuable and satisfactory for all the parties.

1. **Keep it short:** Whatever the children are asked to do, make it fit their time span. This will reduce the effect of satisficing by keeping their motivation high. For

young children, five minutes spent in a written survey is generally long enough, more time can be given, as the children get older.

2. **Pilot the language:** In a survey using written language, children will take short cuts if they cannot read the questions. Teachers can be useful in checking to see if the words used in the survey make sense, they may point out where words may mean something different to children. Avoid ambiguity by piloting with sample children.
3. **Provide assistance for non / poor readers:** Even with the language checked, there will be some children who may understand the words but not the questions. Try to read out written questions if possible, doing this for all the children (as some will not admit to not understanding the questions).
4. **Limit the writing:** Children often do not write what they want to say, as they cannot spell the words they want, cannot find the words for things they want to say, or cannot form the letters for the words that they have in mind. Children can be helped by encouraging the drawing of pictures, the use of images and by providing essential words for them to copy.
5. **Use appropriate tools and methods:** Reduce the effects of suggestibility and satisficing by using special methods. The Fun Toolkit provides tools to assist children in discriminating between rival products [22]. In interviews, use visual props to help articulate ideas. If interviewing, consider taping the discussion so that the amount of ‘suggesting’ can be examined later.
6. **Make it fun:** Introduce glue, scissors, sticky tape or coloured pencils to make the experience fun for the children. If at all possible print questions in colour and supply thank you certificates when the children have finished participating
7. **Expect the unexpected:** Have a back up plan. If an entire project depends on the results of a survey with children it may well fail! Triangulate where possible ideas include observations and post hoc reports from researchers and children..
8. **Don’t take it too seriously:** One of the great pitfalls in research and development work is to read too much into data. The information gained from a single group of children in a single place is not likely to be especially generalisable. Avoid the temptation to apply statistical tests to children’s responses, rather look for trends and outliers! It has been noted that in some instances, children’s responses are not very stable over time [33] so it may be that all that can be elicited from a survey is a general feel for a product or a concept.

9. **Be nice:** As outlined earlier, interviewer effects are significant. To get the most from children, interviewers and researchers need to earn the right to talk to them. This may require several visits and may require an investment of time to learn about their culture and their concerns.

There is no doubt that designing and carrying out good surveys takes practise and patience but following these guidelines may avoid many of the common errors and minimise harmful effects.

CONCLUSION

Because a survey is, by definition, designed, it will always be restrictive. Researchers and developers of interactive products are generally not specialists in survey design and so invariably produce questions and suggested answers that are far from perfect. It is common, and not unexpected, to find that in many studies, the questions are asked in such a way that the answers are invariably the ones the survey designers wanted to hear.

Given the inherent difficulties with survey methods for children, and the survey designer's inadequacy, a case could be made for discouraging these methods in Child Computer Interaction. This approach might gain favour with the empiricists but the value of the survey method to the Child Computer Interaction community is not its validity or its generalisability, but rather the opportunity that these methods provide for researchers and designers to interact with children, to gather their language, and to value their differences.

Perhaps success in a survey in Child Computer Interaction is not to do with stability of responses or reliability of reports but is measured by the answers to two questions for the survey designer, these being: 'Did I learn anything useful? Did I do anything useful?'

It is a privilege to be able to carry out design and evaluation surveys with children. Researchers and developers get to see into the children's worlds and get to glimpse at their dreams and ideals. This requires care and concern, in the words of WB Yeats, "*Tread softly because (you) we tread on (my) their dreams*". It is especially important to neither waste the children's time nor ride roughshod over their opinions.

The guidelines presented in this paper are intended to assist practitioners to carry out careful and gently executed surveys that respect the children and protect their ideals. Much of the literature pertaining to surveying children focuses on what not to do and on the precautions that need to be taken to safeguard the data, future research by the authors will focus on the cost / benefits of surveys in CCI and on refining the methods that need to be taken to provide a special experience for the children.

Acknowledgements

Hundreds of children have freely shared their opinions and views with the authors. They are acknowledged here for their willingness to provide windows into their worlds, for moments of amusement, and for the many special insights they have provided.

REFERENCES

1. Airey, S., et al. Rating Children's Enjoyment of Toys, Games and Media. Procs 3rd World Congress of International Toy Research on Toys, Games and Media. (London. 2002).
2. Bark, I., A. Folstad, and J. Gulliksen. Use and Usefulness of HCI Methods: Results from an exploratory Study among Nordic HCI Practitioners. Procs HCI 2005. (Edinburgh, UK: Springer. 2005) 201 - 217.
3. Bogdan, R.C. and S.K. Biklen, Qualitative research for education: An introduction to theory and methods. 1st ed., Boston: Allyn and Bacon. (1998)
4. Borgers, N. and J. Hox, Item Non response in Questionnaire Research with Children. Journal of Official Statistics, 17(2): 2001. 321 - 335.
5. Borgers, N., J. Hox, and D. Sikkel, Response Effects in Surveys on Children and Adolescents: The Effect of Number of Response Options, Negative Wording, and Neutral Mid-Point. Quality and Quantity, 38(1): 2004. 17 - 33.
6. Borgers, N., J. Hox, and D. Sikkel, Response Quality in Research with Children and Adolescents: The Effect of Labelled Response Opinions and Vague Quantifiers. International Journal of Public Opinion Research, 15(1): 2002. 83 - 94.
7. Borgers, N., E.D. Leeuw, and J. Hox, Children as Respondents in Survey Research: Cognitive Development and Response Quality. Bulletin de Methodologie Sociologique, 66: 2000. 60 - 75.
8. Breakwell, G., Research Methods in Psychology. London: SAGE Publications. (1995)
9. Bruck, M., S.J. Ceci, and L. Melnyk, External and Internal Sources of Variation in the Creation of False Reports in Children. Learning and Individual Differences, 9(4): 1997. 269 - 316.
10. Greig, A. and A. Taylor, Doing Research with Children. London: Sage. (1999)
11. Hanna, E., et al., The Role of Usability Research in Designing Children's Computer Products, in The Design of Children's Technology, A. Druin, Editor. Morgan Kaufmann: San Francisco. (1999). 4-26.
12. Hanna, L., D. Neapolitan, and K. Ridsen. Evaluating Computer Game Concepts with Children. Procs IDC2004. (Maryland, US: ACM Press. 2004) 49 - 56.
13. Holoday, B. and A. Turner-Henson, Response effects in surveys with school-age children. Nursing Research (methodological corner), 38(248 - 250): 1989.

14. Krosnick, J.A., Response Strategies for coping with the Cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5: 1991. 213 - 236.
15. MacFarlane, S.J., G. Sim, and M. Horton. Assessing Usability and Fun in Educational Software. *Procs IDC2005*. (Boulder, Co: ACM Press. 2005).
16. Metaxas, G., et al. SCORPIODROME: An exploration in Mixed Reality Social Gaming for Children. *Procs ACE 2005*. (Valencia, Spain: ACM Press. 2005).
17. Newman, K., Albert in Africa: Online Role-Playing and Lessons from Improvisational Theatre. *Computers in Entertainment*, 3(3): 2005. 4.
18. Peterson, C. and M. Bell, Children's memory for traumatic injury. *Child Development*, 67(6): 1996. 3045 - 70.
19. Read, J.C., The ABC of CCI. *Interfaces*, 62: 2005. 8 - 9.
20. Read, J.C. Optimising Text Input for Young Children using Computers for Creative Writing Tasks. *Procs HCI 2002*. (London, England. 2002).
21. Read, J.C., et al. An Investigation of Participatory Design with Children - Informant, Balanced and Facilitated Design. *Procs Interaction Design and Children*. (Eindhoven: Shaker Publishing. 2002) 53 - 64.
22. Read, J.C., S.J. MacFarlane, and C. Casey. Endurability, Engagement and Expectations: Measuring Children's Fun. *Procs Interaction Design and Children*. (Eindhoven: Shaker Publishing. 2002) 189 - 198.
23. Read, J.C., S.J. MacFarlane, and C. Casey. Expectations and Endurability - Measuring Fun. *Procs Computers and Fun 4*. (York, England. 2001).
24. Read, J.C., S.J. MacFarlane, and C. Casey. Measuring the Usability of Text Input Methods for Children. *Procs HCI2001*. (Lille, France: Springer Verlag. 2001) 559 - 572.
25. Read, J.C., S.J. MacFarlane, and C. Casey. What's going on? Discovering what Children understand about Handwriting Recognition Interfaces. *Procs IDC 2003*. (Preston, UK: ACM Press. 2003) 135 - 140.
26. Read, J.C., S.J. MacFarlane, and A.G. Gregory. Requirements for the Design of a Handwriting Recognition Based Writing interface for Children. *Procs Interaction Design and Children 2004*. (Maryland, US: ACM Press. 2004).
27. Read, J.C., S.J. MacFarlane, and M. Horton. The Usability of Handwriting Recognition for Writing in the Primary Classroom. *Procs HCI 2004*. (Leeds, UK: Springer Verlag. 2004) 135 - 150.
28. Robertson, J. and J. Good. Children's narrative development through computer game authoring. *Procs IDC2004*. (Maryland, US: ACM Press. 2004) 57 - 64.
29. Said, N. An Engaging Multimedia Design Model. *Procs IDC2004*. (Maryland, US: ACM Press. 2004) 169 -172.
30. Scullin, M.H. and S.J. Ceci, A Suggestibility Scale for Children. *Personality and Individual Differences*, 30: 2001. 843 - 856.
31. Shields, B.J., et al., Predictors of a child's ability to use a visual analogue scale. *Child: Care, Health and Development*, 29(4): 2003. 281 - 290.
32. Tobey, A. and G. Goodman, Children's eyewitness memory: effects of participation and forensic context. *Child Abuse and Neglect*, 16: 1992. 807 - 821.
33. Vaillancourt, P.M., Stability of children's survey responses. *Public opinion quarterly*, 37: 1973. 373 - 387.
34. Williams, M., O. Jones, and C. Fleuriot. Wearable Computing and the geographies of urban childhood - working with children to explore the potential of new technology. *Procs IDC2003*. (Preston, UK: ACM Press. 2003) 111 - 118.